

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-018085

(43)Date of publication of application : 22.01.1999

(51)Int.Cl.

H04N 7/24
H04N 7/32
// H03M 7/36

(21)Application number : 10-157181

(71)Applicant : GENERAL INSTR CORP

(22)Date of filing : 05.06.1998

(72)Inventor : CHEN XUEMIN
LUTHRA AJAY
RAJAN GANESH
NARASIMHAN MANDAYAM

(30)Priority

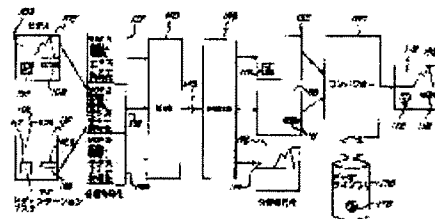
Priority number : 97 869493 Priority date : 05.06.1997 Priority country : US

(54) TEMPORALLY AND SPATIALLY SCALABLE ENCODING FOR VIDEO OBJECT PLANE

(57)Abstract:

PROBLEM TO BE SOLVED: To impart a temporal and spatial scalability function for encoding a video signal including a video object plane(VOP) by differentially encoding an up sample VOP, which is provided from the VOP of an input video sequence, while using the VOP of the input video sequence.

SOLUTION: VOP data from frames 105 and 115 are supplied for separating an encoding function. Especially, VOP 117-119 respectively perform shaping, moving and texture encoding at encoders 137-139. One specified piece of pixel data in the VOP of the input video sequence are down-sampled, and a base layer VOP having a decreased spatial resolution is given. The pixel data in the base layer VOP at one part at least are up-sampled, and an up-sampling VOP is given in a reinforced later. The up-sampling VOP is differentially encoded, while using specified one VOP of the input video sequence.



(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-18085

(43)公開日 平成11年(1999) 1月22日

(51)Int.Cl.*

識別記号

F I

H 0 4 N 7/24

H 0 4 N 7/13

Z

7/32

H 0 3 M 7/36

// H 0 3 M 7/36

H 0 4 N 7/137

Z

審査請求 未請求 請求項の数45 O L (全 20 頁)

(21)出願番号 特願平10-157181

(71)出願人 598045380

(22)出願日 平成10年(1998) 6月5日

ジェネラル・インストルメント・コーポ
レイション

(31)優先権主張番号 8 6 9 4 9 3

アメリカ合衆国ペンシルベニア州ホース
ム, トーナメント・ドライブ 101

(32)優先日 1997年6月5日

(72)発明者 スーミン・チェン

(33)優先権主張国 米国 (US)

アメリカ合衆国カリフォルニア州サンディ
エゴ, フォックスクロフト・プレイス
8560

(72)発明者 アジャイ・ルスラ

アメリカ合衆国カリフォルニア州サンディ
エゴ, スティバ・コート 8714

(74)代理人 弁理士 竹内 澄夫 (外1名)

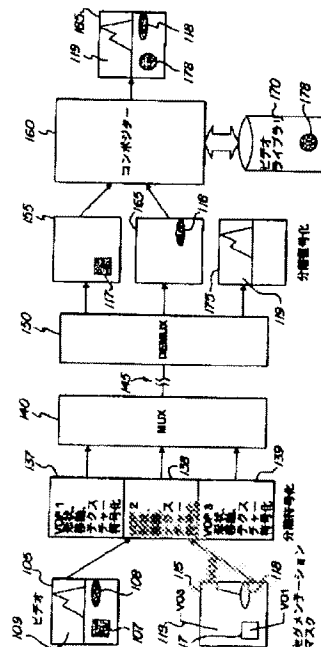
最終頁に続く

(54)【発明の名称】 ビデオオブジェクト平面のための時間的及び空間的スケーラブル符号化

(57)【要約】 (修正有)

【課題】 MPEG-4規格で使用されるようなビデオオブジェクト平面 (VOPs) を含むビデオ信号の符号化の効果的機能を与える。

【解決手段】 入力ビデオシーケンス内にビデオオブジェクト平面を含むスケーリングされたフィールドモードビデオを適宜圧縮することによって符号化効率が改善される。強化レイヤー内のアップサンプリングされたVOPは線形基準に基づいた入力ビデオシーケンスに対しより大きな相関関係を与えるべく並べ替えられる。生成された残差はDCTのような空間変換を使って符号化される。ベースレイヤーVOPに対してすでに決定された移動ベクトルをスケーリングすることによって強化レイヤーVOPを符号化するために、移動補償手法を用いる。。その中心がスケーリングされた移動ベクトルによって定義されるところの縮小した調査範囲が与えられる。移動補償手法はスケーリングされたフレームモード若しくはフィールドモードビデオとともに使用するのに適している。



【特許請求の範囲】

【請求項1】対応するベースレイヤー及び強化レイヤー内で通信するためにビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスをスケーリングする方法であり、前記入カビデオシーケンス内の前記VOPが付随する空間解像度及び時間解像度を有する方法であって、減少した空間解像度を有する第1ベースレイヤーVOPを与えるべく、前記入カビデオシーケンスの前記VOPの第1の特定のひとつの画素データをダウンサンプリングする工程と、前記強化レイヤー内に第1アップサンプルVOPを与えるべく前記第1ベースレイヤーVOPの少なくとも一部の画素データをアップサンプリングする工程と、前記第1ベースレイヤーVOPに対応する時間的位置において前記強化レイヤー内で通信するために、前記入カビデオシーケンスの前記VOPの前記第1の特定のひとつを使って前記第1アップサンプルVOPを差分符号化する工程と、から成る方法。

【請求項2】前記入カビデオシーケンス内の前記VOPはフィールドモードVOPであるところの請求項1に記載の方法であって、前記差分符号化工程はさらに、もし画素データのラインが並べ替え基準に一致すれば、前記第1アップサンプルVOPの前記画素データのの前記ラインをフィールドモードで並べ替える工程と、前記第1アップサンプルVOPの画素データと前記入カビデオシーケンスの前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定する工程と、変換係数を与えるべく前記残差を空間的に変換する工程と、から成るところの方法。

【請求項3】請求項2に記載の方法であって、反対フィールドのラインのルミナンス値の差の和が同一フィールドのラインのルミナンスデータの差の和及びバイアス項よりも大きいとき、前記第1アップサンプルVOPの画素データのの前記ラインは前記並べ替え基準に一致する、ところの方法。

【請求項4】請求項1から3に記載の方法であって、さらに減少した空間解像度を有する第2ベースレイヤーVOPを与えるべく、前記入カビデオシーケンスの前記VOPの第2の特定のひとつの画素データをダウンサンプリングする工程と、前記第1アップサンプルVOPに対応する第2アップサンプルVOPを前記強化レイヤー内に与えるべく、前記第2ベースレイヤーVOPの少なくとも一部の画素データをアップサンプリングする工程と、前記第1及び第2アップサンプルVOPに対応する中間VOPを予測するべく、前記第1及び第2ベースレイヤーVOPの少なくとも一つを使用する工程と、前記第1及び第2アップサンプルVOPの中間の時間的位置において前記強化レイヤー内で通信するために前記中間VOPを符号化する工程と、から成る方法。

【請求項5】請求項4に記載の方法であって、前記強化レイヤーは前記ベースレイヤーより高い時間解像度を有

し、前記ベース及び強化レイヤーは、(a)ピクチャー・イン・ピクチャー(PIP)イメージが前記ベースレイヤー内で運ばれるところのPIPレイバリティ、及び(b)プレビュー・アクセス・イメージが前記ベースレイヤー内で運ばれるところのプレビュー・アクセス・チャンネルレイバリティ、の少なくともひとつを与えるべく適応される、ところの方法。

【請求項6】請求項1から5のいずれかに記載の方法であって、前記ベースレイヤーは優先順位がより高くかつビット速度がより小さいデータを選ぶべく適応され、前記強化レイヤーは優先順位がより低くビット速度がより大きいデータを選ぶべく適応される、ところの方法。

【請求項7】対応するベースレイヤー及び強化レイヤー内で通信するためにビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスをスケーリングする方法であり、前記入カビデオシーケンス内の前記VOPが付随する空間解像度及び時間解像度を有する方法であって、第1ベースレイヤーVOPとして前記ベースレイヤー内で通信するために、前記入カビデオシーケンスの前記VOPの第1の特定のひとつを与える工程と、前記第1ベースレイヤーVOPに対応する時間的位置において第1ダウンサンプルVOPとして前記強化レイヤー内で通信するために、前記第1ベースレイヤーVOPの少なくとも一部の画素データをダウンサンプリングする工程と、比較VOPを与えるべく前記VOPの前記第1の特定のひとつの対応する画素データをダウンサンプリングする工程と、前記比較VOPを使って前記第1ダウンサンプルVOPを差分符号化する工程と、から成る方法。

【請求項8】請求項7に記載の方法であって、さらに前記第1ベースレイヤーVOPの画素データと前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定し、変換係数を与えるべく前記残差を空間変換することによって、前記VOPの前記第1の特定のひとつを使って前記第1ベースレイヤーVOPを差分符号化する工程、から成る方法。

【請求項9】請求項8に記載の方法であって、前記入カビデオシーケンス内の前記VOPはフィールドモードVOPであり、前記第1ベースレイヤーVOPは、もし画素データのラインが並べ替え基準に一致すれば、前記決定工程に先だって前記第1ベースレイヤーVOPの前記画素データのラインをフィールドモードで並べ替える工程によって差分符号化される、ところの方法。

【請求項10】請求項9に記載の方法であって、反対フィールドのラインのルミナンス値の差の和が、同一フィールドのラインのルミナンスデータの差の和及びバイアス項よりも大きいとき、前記第1ベースレイヤーVOPの画素データのの前記ラインは前記並べ替え基準に一致する、ところの方法。

【請求項11】請求項7から10のいずれかに記載の方法であって、さらに第2ベースレイヤーVOPとして前記

ベースレイヤー内で通信するために、前記入カビデオシーケンスの前記VOPの第2の特定のひとつを与える工程と、前記第2ベースレイヤーVOPに対応する時間的位置において第2ダウンサンプルVOPとして前記強化レイヤー内で通信するために、前記第2ベースレイヤーVOPの少なくとも一部の画素データをダウンサンプリングする工程と、比較VOPを与えるべく前記VOPの前記第2の特定のひとつの対応する画素データをダウンサンプリングする工程と、前記第2ダウンサンプルVOPを前記比較VOPを使って差分符号化する工程と、前記第1及び第2ダウンサンプルVOPに対応する中間VOPを予測するべく、前記第1及び第2ベースレイヤーVOPの少なくとも一つを使用する工程と、前記第1及び第2アップサンプルVOPの中間の時間的位置において前記強化レイヤー内で通信するために前記中間VOPを符号化する工程と、から成る方法。

【請求項12】請求項7から11のいずれかに記載の方法であって、ベース及び強化レイヤーは、強化レイヤー内のイメージデータがベースレイヤー内のイメージデータより低い空間解像度を有するところの立体ビデオケイパリティを与えるべく適応される、ところの方法。

【請求項13】双方向予測されたビデオオブジェクト平面(B-VOP)を符号化するための方法であって、対応するベースレイヤー及び強化レイヤー内で通信するために、ビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスをスケーリングする工程と、前記入カビデオシーケンスVOPに対応する第1及び第2ベースレイヤーVOPを前記ベースレイヤー内に与える工程と、前記第2ベースレイヤーVOPが移動ベクトルMV_pに従って前記第1ベースレイヤーVOPから予測される工程と、前記第1及び第2ベースレイヤーVOPの中間の時間的位置において前記強化レイヤー内に前記B-VOPを与える工程と、前記移動ベクトルMV_pをスケーリングすることによって得られる、(a)前方移動ベクトルMV_f、及び(b)後方移動ベクトルMV_bの少なくとも一つを使って前記B-VOPを符号化する工程と、から成る方法。

【請求項14】請求項13に記載の方法であって、時間的距離TR_pは前記第1及び第2ベースレイヤーVOPを分離し、時間的距離TR_bは前記第1ベースレイヤーVOP及び前記B-VOPを分離し、m/nは、B-VOPの空間解像度に対する第1及び第2ベースレイヤーVOPの空間解像度の比であり、(a)前記前方移動ベクトルMV_fは関係 $MV_f = (m/n) \cdot TR_b \cdot MV_p / TR_p$ にしたがって決定されるか、(b)前記後方移動ベクトルMV_bは関係 $MV_b = (m/n) \cdot (TR_b - TR_p) \cdot MV_p / TR_p$ にしたがって決定されるか、の少なくともひとつである、ところの方法。

【請求項15】請求項13または14に記載の方法であって、さらに(a)その中心が前記前方移動ベクトルMV_fに従って決定されるところの前記第1ベースレイヤーVOPの調査範囲、及び(b)その中心が前記後方移動ベクトルM

V_bに従って決定されるところの前記第2ベースレイヤーVOPの調査範囲、の少なくともひとつを使って前記B-VOPを符号化する工程、から成る方法。

【請求項16】スケーリングされ、対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するための方法であり、前記入カビデオシーケンス内の前記VOPは付随する空間解像度及び時間解像度を有し、前記入カビデオシーケンスの前記VOPの第1の特定のひとつの画素データが、ダウンサンプリングされ、減少した空間解像度を有する第1ベースレイヤーVOPとして運ばれ、前記第1ベースレイヤーVOPの少なくとも一部の画素データが、アップサンプリングされ、前記第1ベースレイヤーVOPに対応する時間的位置において前記強化レイヤー内で第1アップサンプルVOPとして運ばれ、前記第1アップサンプルVOPは、前記入カビデオシーケンスの前記VOPの前記第1の特定のひとつを使って差分符号化される、ところの方法であって、当該方法が前記付随する空間解像度を復元するべく前記第1ベースレイヤーVOPの前記画素データをアップサンプリングする工程と、前記付随する空間解像度を有する出力ビデオ信号を与えるべく、前記第1アップサンプルVOP、及び前記復元された付随する空間解像度を有する前記第1ベースレイヤーVOPを処理する工程と、から成る方法。

【請求項17】請求項16に記載の方法であって、前記入カビデオシーケンス内の前記VOPはフィールドモードVOPであり、前記第1アップサンプルVOPは、もし前記画素データのラインが並べ替え基準に一致すれば、前記第1アップサンプルVOPの前記画素データのラインをフィールドモードで並べ替え、その後前記第1アップサンプルVOPの画素データと前記入カビデオシーケンスの前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定し、かつ変換係数を与えるべく前記残差を空間変換することによって、差分符号化される、ところの方法。

【請求項18】請求項17に記載の方法であって、反対フィールドのラインのルミナンス値の差の和が同一フィールドのラインのルミナンスデータの差の和及びバイアス項より大きいとき、前記第1アップサンプルVOPの画素データの前記ラインは前記並べ替え基準に一致する、ところの方法。

【請求項19】請求項16から18のいずれかに記載の方法であって、前記入カビデオシーケンスの前記VOPの第2の特定のひとつが、減少した空間解像度を有する第2のベースレイヤーVOPを与えるべくダウンサンプリングされ、前記第2ベースレイヤーVOPの少なくとも一部の画素データが、前記第1アップサンプルVOPに対応する第2アップサンプルVOPを前記強化レイヤー内に与えるべくアップサンプリングされ、前記第1及び第2のベースレイヤーVOPの少なくとも一つが、前記第1及び第

2 アップサンプルVOPに対応する中間VOPを予測するべく使用され、前記中間VOPは、前記第1及び第2アップサンプルVOPの中間の時間的位置において前記強化レイヤー内で通信するために符号化される、ところの方法。

【請求項20】請求項19に記載の方法であって、前記強化レイヤーは前記ベースレイヤーよりも高い時間解像度を有し、前記ベース及び強化レイヤーは、(a)ピクチャー・イン・ピクチャー(PIP)イメージが前記ベースレイヤー内で運ばれるところのPIPキーパリティ、及び(b)プレビューアクセスイメージが前記ベースレイヤー内で運ばれるところのプレビュー・アクセス・チャンネルキーパリティ、の少なくともひとつを与えるべく適応される、ところの方法。

【請求項21】請求項16から20のいずれかに記載の方法であって、前記ベースレイヤーは優先順位がより高くかつビット速度のより小さいデータを運ぶべく適応され、前記強化レイヤーは優先順位がより低くかつビット速度がより大きいデータを運ぶべく適応される、ところの方法。

【請求項22】スケーリングされ、対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するための方法であり、前記入力ビデオシーケンス内の前記VOPは付随する空間解像度及び時間解像度を有し、前記入力ビデオシーケンスの前記VOPの第1の特定のひとつが、第1ベースレイヤーVOPとして前記ベースレイヤー内に与えられ、前記第1ベースレイヤーVOPの少なくとも一部の画素データが、ダウンサンプリングされ、かつ前記第1ベースレイヤーVOPに対応する時間的位置において第1ダウンサンプルVOPとして前記強化レイヤー内で運ばれ、前記VOPの前記第1の特定のひとつの対応する画素データが比較VOPを与えるべくダウンサンプリングされ、前記第1ダウンサンプルVOPは前記比較VOPを使って差分符号化される、ところの方法であって、当該方法が、前記付随する空間解像度を復元するべく前記第1ダウンサンプルVOPの前記画素データをアップサンプリングする工程と、前記付随する空間解像度を有する出力ビデオ信号を与えるよう、前記復元された付随する空間解像度を有する前記第1強化レイヤーVOP及び前記第1ベースレイヤーVOPを処理する工程と、から成る方法。

【請求項23】請求項22に記載の方法であって、前記第1ベースレイヤーVOPは、前記第1ベースレイヤーVOPの画素データと前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定し、変換係数を与えるべく該残差を空間変換することによって、前記VOPの前記第1の特定のひとつを使って差分符号化される、ところの方法。

【請求項24】請求項23に記載の方法であって、前記入力ビデオシーケンス内の前記VOPはフィールドモードVOPであり、また前記第1ベースレイヤーVOPは、もし画

素データのラインが並べ替え基準に一致すれば、前記残差を決定するに先だって前記第1ベースレイヤーVOPの前記画素データのラインをフィールドモードで並べ替えることによって差分符号化される、ところの方法。

【請求項25】請求項24に記載の方法であって、反対フィールドのラインのルミナンス値の差の和が同一フィールドのラインのルミナンスデータの差の和及びバイアス項より大きいとき、前記第1ベースレイヤーVOPの画素データの前記ラインは前記並べ替え基準に一致する、ところの方法。

【請求項26】請求項22から25のいずれかに記載の方法であって、前記入力ビデオシーケンスの前記VOPの第2の特定のひとつが第2ベースレイヤーVOPとして前記ベースレイヤー内に与えられ、前記第2ベースレイヤーVOPの少なくとも一部の画素データがダウンサンプリングされ、かつ前記第2ベースレイヤーVOPに対応する時間的位置において第2ダウンサンプルVOPとして前記強化レイヤー内で運ばれ、前記VOPの前記第2の特定のひとつの対応する画素データが比較VOPを与えるべくダウンサンプリングされ、前記第2ダウンサンプルVOPは前記比較VOPを使って差分符号化され、前記第1及び第2ベースレイヤーVOPの少なくともひとつは前記第1及び第2アップサンプルVOPに対応する中間VOPを予測するために使用され、前記中間VOPは前記第1及び第2アップサンプルVOPの間である時間的位置において前記強化レイヤー内で通信するために符号化される、ところの方法。

【請求項27】請求項22から26のいずれかに記載の方法であって、前記ベース及び強化レイヤーは、前記強化レイヤー内のイメージデータが前記ベースレイヤー内のイメージデータより低い空間解像度を有するところの立体ビデオキーパリティを与えるべく適応される、ところの方法。

【請求項28】スケーリングされ、データストリーム内の対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するための方法であり、前記入力ビデオシーケンス内の前記VOPは付随する空間解像度及び時間解像度を有し、前記入力ビデオシーケンスVOPに対応する第1及び第2ベースレイヤーVOPが前記ベースレイヤー内に与えられ、前記第2ベースレイヤーVOPは、移動ベクトル MV_0 に従って前記第1ベースレイヤーVOPから予測され、双方向予測されたビデオオブジェクト平面(B-VOP)が、前記第1及び第2ベースレイヤーVOPの中間の時間的位置において前記強化レイヤー内に与えられ、前記B-VOPは前記移動ベクトル MV_0 をスケーリングすることによって得られる前方移動ベクトル MV_f 及び後方移動ベクトル MV_b を使って符号化される、ところの方法であって、当該方法が前記前方移動ベクトル MV_f 及び前記後方移動ベクトル MV_b を前記データストリームから回復する

工程と、前記前方移動ベクトル MV_f 及び前記後方移動ベクトル MV_b を使って前記B-VOPを復元化する工程と、から成る方法。

【請求項29】請求項28に記載の方法であって、時間的距離 TR_b は前記第1及び第2ベースレイヤーVOPを分離し、時間的距離 TR_b は前記第1ベースレイヤーVOP及び前記B-VOPを分離し、 m/n は、B-VOPの空間解像度に対する第1及び第2ベースレイヤーVOPの空間解像度の比であり、(a)前記前方移動ベクトル MV_f は関係 $MV_f=(m/n) \cdot TR_b \cdot MV_p/TR_p$ にしたがって決定されるか、(b)前記後方移動ベクトル MV_b は関係 $MV_b=(m/n) \cdot (TR_b-TR_p) \cdot MV_p/TR_p$ にしたがって決定されるか、の少なくともひとつである、ところの方法。

【請求項30】請求項28または29に記載の方法であって、前記B-VOPは、(a)その中心が前記前方移動ベクトル MV_f に従って決定されるところの前記第1ベースレイヤーVOPの調査範囲、及び(b)その中心が前記後方移動ベクトル MV_b に従って決定されるところの前記第2ベースレイヤーVOPの調査範囲、の少なくともひとつを使って符号化される、ところの方法。

【請求項31】スケーリングされ、対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するためのデコーダ装置であり、前記入力ビデオシーケンス内の前記VOPは付随する空間解像度及び時間解像度を有し、前記入力ビデオシーケンスの前記VOPの第1の特定のひとつの画素データは、ダウンサンプリングされかつ減少した空間解像度を有する第1ベースレイヤーVOPとして運ばれ、前記第1ベースレイヤーVOPの少なくとも一部の画素データは、アップサンプリングされかつ前記第1ベースレイヤーVOPに対応する時間的位置において前記強化レイヤー内で第1アップサンプルVOPとして運ばれ、前記第1アップサンプルVOPは、前記入力ビデオシーケンスの前記VOPの前記第1の特定のひとつを使って差分符号化される、ところの装置であって、当該装置が前記付随する空間解像度を復元するべく前記第1ベースレイヤーVOPの前記画素データをアップサンプリングする手段と、前記付随する空間解像度を有する出力ビデオ信号を与えるよう、前記第1アップサンプルVOP及び前記復元された付随する空間解像度を有する前記第1ベースレイヤーVOPを処理する手段と、から成る装置。

【請求項32】請求項31に記載の装置であって、前記入力ビデオシーケンス内の前記VOPはフィールドモードVOPであり、前記第1アップサンプルVOPは、もし前記画素データのラインが並べ替え基準に一致すれば、前記第1アップサンプルVOPの前記画素データのラインをフィールドモードで並べ替え、その後前記第1アップサンプルVOPの画素データと前記入力ビデオシーケンスの前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定し、また変換係数を与えるべく前記残差

を空間変換することによって、差分符号化される、ところの装置。

【請求項33】請求項31または32に記載の装置であって、反対フィールドのラインのルミナンス値の差の和が同一フィールドのラインのルミナンスデータの差の和及びバイアス項より大きいとき、前記第1アップサンプルVOPの画素データの前記ラインは前記並べ替え基準に一致する、ところの装置。

【請求項34】請求項31から33のいずれかに記載の装置であって、前記入力ビデオシーケンスの前記VOPの第2の特定のひとつが、減少した空間解像度を有する第2ベースレイヤーVOPを与えるべくダウンサンプリングされ、前記第2ベースレイヤーVOPの少なくとも一部の画素データが、前記第1アップサンプルVOPに対応する第2アップサンプルVOPを前記強化レイヤー内に与えるべくアップサンプリングされ、前記第1及び第2のベースレイヤーVOPの少なくともひとつが、前記第1及び第2アップサンプルVOPに対応する中間VOPを予測するのに使用され、前記中間VOPは、前記第1及び第2アップサンプルVOPの中間の時間的位置において前記強化レイヤー内で通信するために符号化される、ところの装置。

【請求項35】請求項34に記載の装置であって、前記強化レイヤーは前記ベースレイヤーよりも高い時間解像度を有し、前記ベース及び強化レイヤーは、(a)ピクチャー・イン・ピクチャー(PIP)イメージが前記ベースレイヤー内で運ばれるところのPIPレイバリティ、及び(b)プレビュー・アクセス・イメージが前記ベースレイヤー内で運ばれるところのプレビュー・アクセス・チャネルレイバリティ、の少なくともひとつを与えるべく適応される、ところの装置。

【請求項36】請求項31から35のいずれかに記載の装置であって、前記ベースレイヤーは優先順位がより高くかつビット速度のより小さいデータを運ぶべく適応され、前記強化レイヤーは優先順位がより低くかつビット速度がより大きいデータを運ぶべく適応される、ところの装置。

【請求項37】スケーリングされ、対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するためのデコーダ装置であり、前記入力ビデオシーケンス内の前記VOPは付随する空間及び時間解像度を有し、前記入力ビデオシーケンスの前記VOPの第1の特定のひとつが、第1ベースレイヤーVOPとして前記ベースレイヤー内に与えられ、前記第1ベースレイヤーVOPの少なくとも一部の画素データがダウンサンプリングされ、かつ前記第1ベースレイヤーVOPに対応する時間的位置において第1ダウンサンプルVOPとして前記強化レイヤー内で運ばれ、前記VOPの前記第1の特定のひとつの対応する画素データが比較VOPを与えるべくダウンサンプリングされ、前記第1ダウンサンプルVOPは前記比較VOPを使

って差分符号化される。ところの装置であって、当該装置が、前記付随する空間解像度を復元するべく前記第1ダウンサンプルVOPの前記画素データをアップサンプリングする手段と、前記付随する空間解像度を有する出力ビデオ信号を与えるよう、前記復元された空間解像度を有する前記第1強化レイヤーVOP及び前記第1ベースレイヤーVOPを処理する手段と、から成る装置。

【請求項38】請求項37に記載の装置であって、前記第1ダウンサンプルVOPは、前記第1ダウンサンプルVOPの画素データと前記入カビデオシーケンスの前記VOPの前記第1の特定のひとつの画素データとの間の差に従う残差を決定し、変換係数を与えるべく該残差を空間変換することによって、差分符号化される。ところの装置。

【請求項39】請求項38に記載の装置であって、前記入カビデオシーケンス内の前記VOPはフィールドモードVOPであり、また前記第1ベースレイヤーVOPは、もし画素データのラインが並べ替え基準に一致すれば、前記残差を決定するに先だって前記第1ベースレイヤーVOPの前記画素データのラインをフィールドモードで並べ替えることによって、差分符号化される。ところの装置。

【請求項40】請求項39に記載の装置であって、反対フィールドのラインのルミナンス値の差の和が同一フィールドのラインのルミナンスデータの差の和及びバイアス項より大きいとき、前記第1ベースレイヤーVOPの画素データの前記ラインは前記並べ替え基準に一致する。ところの装置。

【請求項41】請求項37から40のいずれかに記載の装置であって、前記入カビデオシーケンスの前記VOPの第2の特定のひとつが第2ベースレイヤーVOPとして前記ベースレイヤー内で通信するために与えられ、前記第2ベースレイヤーVOPの少なくとも一部の画素データが、前記第1アップサンプルVOPに対応する第2ダウンサンプルVOPを前記強化レイヤー内に与えるべくダウンサンプリングされ、前記第1及び第2ベースレイヤーVOPの少なくともひとつが前記第1及び第2ダウンサンプルVOPに対応する中間VOPを予測するのに使用され、前記中間VOPは前記第1及び第2アップサンプルVOPの間である時間的位置において前記強化レイヤー内で通信するために符号化される。ところの装置。

【請求項42】請求項37から41のいずれかに記載の装置であって、前記ベース及び強化レイヤーは、前記強化レイヤー内のイメージデータが前記ベースレイヤー内のイメージデータより低い空間解像度を有するところの立体ビデオケイパビリティを与えるべく適応される。ところの装置。

【請求項43】スケーリングされ、データストリーム内の対応するベースレイヤー及び強化レイヤー内で通信されるビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスを回復するためのデコーダ装置であり、前記入カビデオシーケンス内の前記VOPは付随する空間解像

度及び時間解像度を有し、前記入カビデオシーケンスVOPに対応する第1及び第2ベースレイヤーVOPが前記ベースレイヤー内に与えられ、前記第2ベースレイヤーVOPは、移動ベクトル MV_p に従って前記第1ベースレイヤーVOPから予測され、双方向予測されたビデオオブジェクト平面(B-VOP)が、前記第1及び第2ベースレイヤーVOPの中間の時間的位置において前記強化レイヤー内に与えられ、前記B-VOPは前記移動ベクトル MV_p をスケーリングすることによって得られる前方移動ベクトル MV_f 及び後方移動ベクトル MV_b を使って符号化される。ところの装置であって、当該装置が前記前方移動ベクトル MV_f 及び前記後方移動ベクトル MV_b を前記データストリームから回復する手段と、前記前方移動ベクトル MV_f 及び前記後方移動ベクトル MV_b を使って前記B-VOPを復元化する手段と、から成る装置。

【請求項44】請求項43に記載の方法であって、時間的距離 TR_p は前記第1及び第2ベースレイヤーVOPを分離し、時間的距離 TR_b は前記第1ベースレイヤーVOP及び前記B-VOPを分離し、 m/n は、B-VOPの空間解像度に対する第1及び第2ベースレイヤーVOPの空間解像度の比であり、(a)前記前方移動ベクトル MV_f は関係 $MV_f = (m/n) \cdot TR_b \cdot MV_p / TR_p$ にしたがって決定されるか、(b)前記後方移動ベクトル MV_b は関係 $MV_b = (m/n) \cdot (TR_b - TR_p) \cdot MV_p / TR_p$ にしたがって決定されるか、の少なくともひとつである。ところの装置。

【請求項45】請求項43または44に記載の装置であって、前記B-VOPは、(a)その中心が前記前方移動ベクトル MV_f に従って決定されるところの前記第1ベースレイヤーVOPの調査範囲、及び(b)その中心が前記後方移動ベクトル MV_b に従って決定されるところの前記第2ベースレイヤーVOPの調査範囲、の少なくともひとつを使って符号化される。ところの装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、デジタルビデオシーケンス内のビデオオブジェクト平面を含むビデオイメージの空間的及び時間的スケーリングを与える方法及び装置に関する。特に、スケーリングされたフレームモード若しくはフィールドモードビデオとともに使用するのに適した移動補償手法が与えられる。また、離散コサイン変換(DCT)のような空間的変換を使用するフィールドモードビデオを適宜圧縮するための手法も与えられる。

【0002】

【従来の技術】本発明は特にさまざまなマルチメディアアプリケーションとともに使用するのに適しており、またここに参考文献として組み込む書類「ISO/IEC/JTC1/S C29/WG11 N1642、題名「MPEG-4 Video Verification Model Version 7.0」1997年4月発行」に記載されたMPEG-4 Verification Model (VM) 7.0 規格と互換性がある。さらに本発明は、立体ビデオ符号化、ピクチャー・イン・

ピクチャー、プレビュー・アクセス・チャンネル、及び非同期転送モード(ATM)通信をもたらす。

【0003】MPEG-4は、デジタルオーディオビジュアルデータの通信、アクセス及び操作のためのフレキシブルな枠組及び符号化ツールのオープンセットを与える新規な符号化規格である。該フレキシブルな枠組は、符号化ツールのさまざまな組み合わせ、及びデータベースブラウジング、情報検索及び双方向通信のようなコンピュータ、テレコミュニケーション、及びエンターテイメント（すなわち、TV及び映画）産業により要求される応用に対するそれら符号化ツールのさまざまな組み合わせに対応する機能性をサポートする。

【0004】MPEG-4はマルチメディア環境内でビデオデータの効果的保存、送信及び操作をもたらす標準化されたコアテクノロジーである。MPEG-4は効果的圧縮、オブジェクトスケーラビリティ、空間的及び時間的スケーラビリティ、並びにエラー弾力性を達成する。

【0005】当該MPEG-4ビデオVMコーデック/デコーダ(CODEC)は、移動補償を伴うブロック及びオブジェクトベースのハイブリッドコーデックである。テキストチャーはオーバーラップされたブロック移動補償(overlapped block-motion compensation)を利用して8×8のDCTで符号化される。オブジェクト形状はアルファマップとして表され、内容ベースの算術符号化(CAE)若しくは修正DCTコーデック（いずれも時間的予測を使用する）を使って符号化される。該コーデックはコンピュータグラフィックスで知られるようなスプライト(sprite)を扱うことができる。ウェーブレット(wavelet)及びスプライト符号化のようなその他の符号化方法もまた空間的応用に対し使用されてもよい。

【0006】移動補償されたテキストチャー符号化はビデオ符号化として周知のアプローチである。そのようなアプローチは3つのステージの処理工程としてモデル化される。第1ステージは移動推定及び補償(ME/MC)並びに2次元空間変換を含む単一処理である。ME/MC及び空間変換の目的は、ビデオシーケンス内の時間的及び空間的相関関係を利用することであり、その結果量子化の歪み率性能及び複雑な制約条件の下での符号化エントロピーを最適化する。ME/MCに対する最も通常の技術はブロックマッチングであり、最も通常の空間変換はDCTであった。しかし、任意形状のVOPの境界ブロックのME/MC及びDCTに対して空間的問題が持ち上がった。

【0007】MPEG-2メインプロファイルはMPEG-4規格に先行し、ここに先行文献として組み込む書類「ISO/IEC JTC1/SC29/WG11 N0702、題名「Information Technology - Generic Coding of Moving Pictures and Associated Audio, Recommendation H.262」1994年3月25日発行」に説明されている。2つまたはそれ以上の独立ビットストリーム若しくはレイヤーを与えるようにMPEG-2メインプロファイルへのスケーラビリティ拡張が定義された。各

レイヤーは単一の高品位信号を形成するよう組み合わせられる。例えば、ベースレイヤー(base layer)はより低品位のビデオ信号を与え、一方強化レイヤー(enhancement layer)はベースレイヤーイメージを強化するような付加的情報を与える。

【0008】特に、空間的及び時間的スケーラビリティは異なるビデオ規格若しくはデコーダ性能の間に互換性をもたらす。空間的スケーラビリティに関して、ベースレイヤービデオは入力ビデオシーケンスより低い空間的解像度を有し、その場合強化レイヤーはベースレイヤーの解像度を入力シーケンスレベルまで復元することができる情報を運ぶ。例えば、国際テレコミュニケーションズユニオン-ラジオセクター(ITU-R)601規格(720×576画素の解像度を有する)に対応する入力ビデオシーケンスは、共通インターチェンジフォーマット(CIF)規格(360×288画素の解像度を有する)に対応するベースレイヤー内で運ばれる。この場合の強化レイヤーは、ベースレイヤーのビデオをITU-R601規格に復元するためにデコーダによって使用される情報を運ぶ。変形的に、強化レイヤーは減少した空間解像度を有することもできる。

【0009】時間的スケーラビリティに関して、ベースレイヤーは入力ビデオシーケンスより低い時間解像度（すなわち、フレーム速度）を有し、一方強化レイヤーは欠フレームを運ぶ。デコーダで組み合わせられるとき、オリジナルフレーム速度が復元される。

【0010】

【発明が解決しようとする課題】したがって、MPEG-4規格で使用されるようなビデオオブジェクト平面(VOPs)を含むビデオ信号の符号化用の時間的及び空間的スケーラビリティ機能を与えることが所望される。それが立体ビデオの符号化、ピクチャー・イン・ピクチャー、プレビュー・アクセス・チャンネル、及び非同期転送モード(ATM)通信に対するレイバビリティを有することが所望される。

【0011】さらに、双方向予測VOPs(B-VOPs)の強化レイヤー予測符号化の移動推定に対する調査範囲のサイズが縮小されるように比較的複雑さが小さく、かつ低コストなcodec設計を有することが所望される。また、残差を決定しかつデータを空間的に変換する前に強化レイヤーVOPの画素ラインを適宜並べ替えることによって、ベース及び強化レイヤーヘスケーリングされるインターレースされたビデオ入力信号を効果的に符号化することが所望される。本発明は上記及び他の利点を有する装置を与える。

【0012】

【課題を解決するための手段】本発明に従って、デジタルビデオシーケンス内のビデオオブジェクト平面(VOPs)のようなビデオイメージの時間的及び空間的スケーリングを与える方法及び装置が与えられる。VOPは完全フレーム及び/またはフレームのサブセットから成り、任意

形状を有することができる。付加的に、複数のVOPが一つのフレーム内に与えられるか、若しくはさらに時間的に同時であってもよい。

【0013】ビデオオブジェクト平面(VOPs)から成る入力ビデオシーケンスに対応するベースレイヤー及び強化レイヤー内で通信するためにスケーリングする方法が与えられ、そこではダウンサンプリングされたデータがベースレイヤー内で運ばれる。入力ビデオシーケンス内のVOPは付随する空間解像度及び時間解像度(例えば、フレーム速度)を有する。

【0014】入力ビデオシーケンスのVOPの第1の特定のひとつの画素データはダウンサンプリングされ、減少した空間解像度を有する第1ベースレイヤーVOPを与える。少なくとも一部の第1ベースレイヤーVOPの画素データはアップサンプリングされ、強化レイヤー内に第1アップサンプルVOPを与える。該第1アップサンプルVOPは入力ビデオシーケンスのVOPの第1の特定のひとつを使って差分符号化され、第1ベースレイヤーVOPに対応する時間的位置で強化レイヤー内に与えられる。

【0015】差分符号化は、第1アップサンプルVOPの画素データと、入力ビデオシーケンスのVOPの第1の特定のひとつの画素データとの間の差に対応する残差を決定する工程を含む。該残差は空間的に変換され、例えば、DCTを使って変換係数が与えられる。

【0016】入力ビデオシーケンス内のVOPがフィールドモードのVOPであるとき、差分符号化は、もし画素データのラインが並べ替え基準に一致すれば、残差を決定する前に第1アップサンプルVOPの画素データラインをフィールドモードで並べ替える工程を含む。該基準とは、反対フィールドライン(例えば、奇数から偶数へ及び偶数から奇数へ)のルミナンス値の差の和が、同一フィールドライン(例えば、奇数から奇数及び偶数から偶数)のルミナンスデータの差の和及びバイアス項よりも大きいかどうかである。

【0017】第1ベースレイヤーVOPのアップサンプリングされた画素データは全第1ベースレイヤーVOPのサブセットでもよく、その結果アップサンプリングされない第1ベースレイヤーVOPの残りの部分はアップサンプリングされた画素データより低い空間解像度を有する。

【0018】第2ベースレイヤーVOP及び強化レイヤー内のアップサンプルVOPは同様の方法で与えられる。第1及び第2ベースレイヤーVOPの一つ若しくは両方は、第1及び第2アップサンプルVOPに対応する中間VOPを予測するために使用される。該中間VOPは、時間的に第1と第2アップサンプルVOPとの間で、強化レイヤー内で通信するために符号化される。

【0019】さらに、第1と第2ベースレイヤーVOPの間に中間ベースレイヤーVOPが存在しないとき、強化レイヤーはベースレイヤーより高い時間的解像度を有してもよい。

【0020】特定の応用において、ベース及び強化レイヤーは、ピクチャー・イン・ピクチャー(PIP)イメージがベースレイヤー内で運ばれるところのPIPレイバリティ、若しくはプレビュー・アクセス・イメージがベースレイヤー内で運ばれるところのプレビュー・アクセス・チャンネルレイバリティを与える。そのような応用において、PIPイメージ若しくは無料プレビューイメージが、減少した空間的及び/または時間的解像度を有することが許容される。ATM応用において、より高い優先順位でより低いビット速度のデータがベースレイヤー内に与えられ、一方より低い優先順位でより高いビット速度のデータが強化レイヤー内に与えられる。この場合、ベースレイヤーは保証された帯域幅に割り当てられるが、強化レイヤーデータは時々失われる。

【0021】ダウンサンプリングされたデータがベースレイヤーでなく強化レイヤー内で運ばれるところのビデオオブジェクト平面(VOP)から成る入力ビデオシーケンスをスケーリングするための方法が与えられる。この方法に関して、入力ビデオシーケンスの第1の特定のひとつのVOPは、例えば空間解像度を変化させることなく第1ベースレイヤーVOPとしてベースレイヤー内に与えられる。少なくとも一部の第1ベースレイヤーVOPの画素データはダウンサンプリングされ、第1ベースレイヤーVOPに対応する時間的位置で強化レイヤー内に対応する第1ダウンサンプルVOPが与えられる。第1の特定のひとつのVOPの対応する画素データは比較VOPを与えるべくダウンサンプリングされ、第1ダウンサンプルVOPは該比較VOPを使って差分符号化される。

【0022】該ベース及び強化レイヤーは、強化レイヤー内のイメージデータがベースレイヤー内のイメージデータより低い空間解像度を有するところの立体ビデオレイバリティを与える。

【0023】双方向的に予測されたビデオオブジェクト平面(B-VOP)を符号化するための方法も与えられる。入力ビデオシーケンスVOPに対応する第1及び第2ベースレイヤーVOPがベースレイヤー内に与えられる。第2ベースレイヤーVOPは、移動ベクトル MV_1 に従って第1ベースレイヤーVOPから予測されるP-VOPである。B-VOPは、時間的に第1と第2ベースレイヤーVOPの間で強化レイヤー内に与えられる。

【0024】B-VOPは、移動ベクトル MV_1 をスケーリングすることによって得られる少なくとも一つの前方移動ベクトル MV_1 及び後方移動ベクトル MV_2 を使って符号化される。この有効な符号化技術によって、リファレンスVOP内で独立の網羅的な調査を実行する必要がなくなる。時間的距離 TR_1 は第1及び第2ベースレイヤーVOPを分離し、一方時間的距離 TR_2 は第1ベースレイヤーVOP及びB-VOPを分離する。

【0025】比 m/n は、第1及び第2ベースレイヤーVOPの空間解像度とB-VOPの空間解像度の比を定義したものの

である。すなわち、ベースレイヤーVOPまたは強化レイヤー内のB-VOPのいずれかは、入力ビデオシーケンスのVOPに関して比 m/n だけダウンサンプリングされる。ここでベースまたは強化レイヤーVOPのいずれかが入力ビデオシーケンスと同じ空間解像度を有することが仮定されている。前方移動ベクトル MV_f は $MV_f = (m/n) \cdot TR_g \cdot MV_p / TR_g$ の関係に従って決定され、一方後方移動ベクトル MV_b は、 $MV_b = (m/n) \cdot (TR_g - TR_p) \cdot MV_p / TR_p$ の関係に従って決定される。 m/n は分数値を含む正数である。

【0026】B-VOPは、第1ベースレイヤーVOPの調査範囲でありその中心が前方移動ベクトル MV_f に従って決定されるところの調査範囲と、第2ベースレイヤーVOPの調査範囲でありその中心が後方移動ベクトル MV_b に従って決定されるところの調査範囲とを使って符号化される。

【0027】対応するデコーダ方法及び装置もまた与えられる。

【0028】

【発明の実施の形態】デジタルビデオシーケンス内にビデオオブジェクト平面(VOP)を含むビデオイメージの時間的及び空間的スケーリングを与えるための方法及び装置が与えられる。

【0029】図1は本発明に従うビデオオブジェクト符号化及び復合化処理を示す。フレーム105は、正方形の前景エレメント107、楕円形の前景エレメント108及び背景エレメント109を含む3つの絵図エレメントを有する。フレーム115において、該エレメントは、セグメンテーションマスクを使って指定されたVOPであり、その結果VOP117は正方形の前景エレメント107を表し、VOP118は楕円前景エレメント108を表し、VOP119は背景エレメント109を表す。VOPは任意形状を有することができ、連続のVOPはビデオオブジェクトとして知られる。完全な矩形ビデオフレームもまたVOPとして考えることも可能である。従って、ここで使用される“VOP”の語は、任意及び非任意のイメージ領域形状の両方を示すのに使用される。セグメンテーションマスクは周知の技術を使って得られ、それはITU-R 601ルミナンスデータと類似のフォーマットを有する。各画素はビデオフレーム内のある領域に属するように識別される。

【0030】フレーム105及びフレーム115からのVOPデータは符号化機能を分離するために供給される。特に、VOP117、118及び119はエンコーダ137、138及び139においてそれぞれ形状、移動及びテクスチャ符号化を経験する。形状符号化に関して、バイナリ及びグレイスケール形状情報が符号化される。移動符号化に関して、形状情報はフレーム内の移動推定を使って符号化される。テクスチャ符号化に関して、DCTのような空間変換が実行され、圧縮のために可変長符号化される変換係数が得られる。

【0031】その後、符号化VOPデータがマルチプレク

サ(MUX)140において組み合わせられ、チャンネル145を通じて転送される。変形的に、データは記録媒体に保存されてもよい。独立のVOP117~119が復合化されかつ回復されるように、受信された符号化VOPデータはデマルチプレクサ(DEMUX)150によって分離される。フレーム155、165及び175は、VOP117、118及び119がそれぞれ復合化されかつ回復され、その結果ビデオライブラリ170とのインターフェースであるコンポジター160を使って個別に扱われることを示している。

【0032】該コンポジターは、ユーザがカスタマイズされたイメージをもたらすよう受信データを編集することができるように、ユーザの家庭に配置されたパーソナルコンピュータのような装置である。例えば、ユーザのパーソナルビデオライブラリ170は受信VOPとは異なる予め保存されたVOP178(例えば、円形)を含むことができる。ユーザは円形VOP178が正方形VOP117と入れ替わるところのフレーム185を組み立ててもよい。その結果フレーム185は受信VOP118及び119並びに局所的に保存されたVOP178を含む。

【0033】他の実施例において、背景VOP119はユーザの選択による背景と置換されてもよい。例えば、テレビニュース放送を観ているとき、アナウンサーはニューススタジオのような背景から分離されたVOPとして符号化され得る。該ユーザはライブラリ170から、またはストック価格若しくは天気情報を有するチャンネルのような他のテレビ番組から背景を選択してもよい。したがって、ユーザはビデオエディターのように行動することができる。

【0034】該ビデオライブラリ170はチャンネル145を通じて受信されるVOPを保存することもでき、またインターネットのようなネットワークを通じてVOP及び他のイメージエレメントにアクセスすることもできる。

【0035】フレーム105がVOPではなくしたがって個別に扱われない領域を含むことができるということを認識すべきである。さらに、フレーム105はVOPを有する必要はない。概して、ビデオセッションは単一のVOP若しくは連続のVOPから成る。

【0036】図1のビデオオブジェクト符号化及び復合化処理によって、パーソナルコンピュータゲーム、仮想環境、グラフィカル・ユーザー・インターフェース、ビデオ会議、インターネットアプリケーション等を含む多くのエンターテイメント、ビジネス及び教育的応用が可能となる。特に、本発明に従うVOPの空間的及び時間的スケーリングに対するケイパビリティがさらに大きなケイパビリティをもたらす。

【0037】図2は本発明に従うビデオオブジェクトコード及びデコーダのブロック図である。図1に略示されたエレメント137~139に対応するエンコーダ201は、入力ビデオデータシーケンス“in”を受信するスケーラビリティ・プレプロセッサ205を含む。強化レイヤーより

も低い空間解像度を有するベースレイヤーの有する空間スケーラビリティを達成するために、信号“in_0”を得るべく“in”が空間的にダウンサンプリングされ、続いてそれは経路217を通じてベースレイヤーエンコーダ220へ与えられる。“in_0”はベースレイヤーエンコーダ220で符号化され、また符号化データはマルチプレクサ(MUX)230へ与えられる。ここでMPEG-4システム及び記述語(MSDL)MUXが使用されてもよい。

【0038】入力ビデオシーケンス“in”がフィールド(インターレース)モードであるとき、ダウンサンプリングは画素データを偶数及び奇数フィールド内に保存しないため、ダウンサンプリングされた信号“in_0”はフレーム(プログレッシブ)モードである。もちろん、“in”がフレームモードであるとき、“in_0”もまたフレームモードである。

【0039】図3に関して以下に詳細に記載されるように、再構成されたイメージデータは経路218を通じてベースレイヤーエンコーダ220から画素のアップサンプリングを実行するミッドプロセッサへ与えられる。フレームモードのアップサンプリングされたイメージデータはその後経路212を通じて強化レイヤーエンコーダ210へ与えられ、そこでそれは経路207を通じてプレプロセッサ205からエンコーダ210へ与えられる入力イメージデータ“in_1”を使って差分的に符号化される。特に、アップサンプリングされた画素データ(例えば、ルミナンスデータ)は残差を得るべく入力イメージデータから減算され、その後それはDCT若しくは他の空間変換を使って符号化される。

【0040】本発明に従って、入力ビデオシーケンスがフィールドモードであるとき、符号化効率は入力ビデオシーケンスのオリジナルの偶数(トップ)及び奇数(ボトム)フィールドに対応するアップサンプリングされた強化レイヤーイメージの画素ラインをグループ化することによって改善される。しばしばフィールド内の画素データは、反対フィールド内のデータに対するよりも同一フィールド内の他の画素データに対して、より大きな相関関係を有するため、このことはある場合に残差の大きさを減少させることができる。したがって、残差の大きさを減少させることによって、イメージデータを符号化するのに必要なビットをより少なくできる。これは以下の図6及びそれに関連する議論によって、さらに明白となる。

【0041】強化レイヤー内のアップサンプリングされたイメージの符号化された残差はMUX230に与えられ、ベ

ースレイヤーデータとともに通信チャンネル245を通じて送信される。変形的に該データは局所的に保存されてもよい。MUX230、チャンネル245、及びDEMUX250はそれぞれ図1のエLEMENT140、145及び150に対応する。

【0042】ベースレイヤーエンコーダ220からミッドプロセッサ215へ与えられたイメージデータは完全フレームVOP若しくは全イメージのサブセットであるVOPのような全ビデオイメージである。さらに、複数のVOPがミッドプロセッサ215へ与えられてもよい。現在MPEG-4は256個のVOPまでサポートする。

【0043】デコーダ299において、符号化されたデータはMPEG-4 MSDL DEMUXのようなデマルチプレクサ(DEMUX)250で受信される。この例においてベースレイヤーデータよりも高い空間解像度を有する強化レイヤーデータは強化レイヤーデコーダ260へ与えられる。ベースレイヤーデータはベースレイヤーデコーダ270に与えられ、そこで信号“out_0”が回復され、経路267を通じてミッドプロセッサ265へ及び経路277を通じてスケラビリティ・ポストプロセッサ280へ与えられる。該ミッドプロセッサは、完全解像度イメージを回復するべくベースレイヤーデータをアップサンプリングすることによって、エンコーダ側のミッドプロセッサ215と同様の方法で動作する。このイメージは強化レイヤーデータ信号“out_1”を回復する際に使用するために経路262を通じて強化レイヤーデコーダ260へ与えられ、その後経路272を通じてスケラビリティ・ポストプロセッサ280へ与えられる。該スケラビリティ・ポストプロセッサ280は、信号“outp_0”として表示するために復合化ベースレイヤーデータを空間的にアップサンプリングする動作を実行し、一方強化レイヤーデータは信号“outp_1”として表示するために出力される。

【0044】エンコーダ201が時間的スケラビリティのために使用されるとき、例えば、ベースレイヤーに対してフレーム速度を減少させるべくプレプロセッサ205は時間的デマルチプレキシング(例えば、ブルダウン処理若しくはフレームドロッピング)を実行する。例えば、フレーム速度を30フレーム/secから15フレーム/secへ減少させるために、すべての他のフレームが落とされる。

【0045】以下の表1はミッドプロセッサ215及び265、スケラビリティ・プレプロセッサ205及びスケラビリティ・ポストプロセッサ280の24個の可能な構成を示したものである。

【表1】

構成	レイヤー	時間解像度	空間解像度	スケーラビリティ プロセッサ	ミッド プロセッサ	スケーラビリティ ポストプロセッサ
1	ベース	Low(High)	Low	ダウンサンプル フィルタリング	アップサンプル フィルタリング	N/C
	強化	Low(High)	High	N/C	N/A	N/C
2	ベース	Low	Low	ダウンサンプル フィルタリング ブルダウン処理	アップサンプル フィルタリング	N/C
	強化	High	High	N/C	N/A	N/C
3	ベース	High	Low	ダウンサンプル フィルタリング	アップサンプル フィルタリング	N/C
	強化	Low	High	ブルダウン処理	N/A	N/C
4	ベース	Low(High)	High	N/C	ダウンサンプル フィルタリング	N/C
	強化	Low(High)	Low	ダウンサンプル フィルタリング	N/A	アップサンプル フィルタリング
5	ベース	Low	High	ブルダウン処理	ダウンサンプル フィルタリング	N/C
	強化	High	Low	ダウンサンプル フィルタリング	N/A	アップサンプル フィルタリング
6	ベース	High	High	N/C	ダウンサンプル フィルタリング	N/C
	強化	Low	Low	ダウンサンプル フィルタリング ブルダウン処理	N/A	アップサンプル フィルタリング
7	ベース	Low(High)	High	N/C	N/C	N/C
	強化	Low(High)	High	N/C	N/A	N/C
8	ベース	Low	High	ブルダウン処理	N/C	N/C
	強化	High	High	N/C	N/A	N/C
9	ベース	High	High	N/C	N/C	N/C
	強化	Low	High	ブルダウン処理	N/A	N/C
10	ベース	Low(High)	Low	ダウンサンプル フィルタリング	N/C	アップサンプル フィルタリング
	強化	Low(High)	Low	ダウンサンプル フィルタリング	N/A	アップサンプル フィルタリング
11	ベース	Low	Low	ダウンサンプル フィルタリング ブルダウン処理	N/C	アップサンプル フィルタリング
	強化	High	Low	ダウンサンプル フィルタリング	N/A	アップサンプル フィルタリング
12	ベース	High	Low	ダウンサンプル フィルタリング	N/C	アップサンプル フィルタリング
	強化	Low	Low	ダウンサンプル フィルタリング ブルダウン処理 a	N/A	アップサンプル フィルタリング

【0046】表1の中で、第1列は構成番号、第2列はレイヤー、第3列は該レイヤーの時間解像度（例えば、highまたはlow）をそれぞれ示している。“Low(High)”がリストされるとき、ベース及び強化レイヤーの時間解像度は両方ともhigh若しくは両方ともlowである。第4列は空間解像度を示す。第5、6及び7列はスケーラビリティ・プレプロセッサ205、ミッドプロセッサ215、及びスケーラビリティ・ポストプロセッサ280の対応する動作を示す。“N/C”は時間若しくは空間解像度に変化

がないことを示し、すなわちノーマルな処理が実行されている。“N/A”は適用不能を意味する。ミッドプロセッサ215、265の動作は強化レイヤーに影響を与えない。

【0047】空間的にスケーリングされた符号化は例として構成1で示される。説明したように、スケール可能コーダ201がVOPの符号化用に使われるとき、プレプロセッサ205は異なる空間解像度を有するVOPの2つのサブストリームを生成する。表1に示されるように、ベースレイヤーは低い空間解像度を有し、強化レイヤーは入力

シーケンスの解像度に対応する高い空間解像度を有する。したがって、ベースレイヤーシーケンス“in_0”は、スケーラビリティ・ブレプロセッサ205において入力ビデオシーケンス“in”をダウンサンプリングすることによって生成される。強化レイヤーシーケンスは、同一の高い空間解像度“in”を達成するべくミッドプロセッサ215、265においてダウンサンプリングされたベースレイヤーシーケンスのアップサンプル・フィルタリングによって生成される。該ポストプロセッサ280はノーマル処理を実行する、すなわちそれは“out_1”若しくは“out_0”の時間若しくは空間解像度を変更しない。

【0048】例えば、ベースレイヤーCIF解像度シーケンス(360×288画素)はITU-R 601解像度入力シーケンス(720×576画素)の2:1ダウンサンプル・フィルタリングより生成される。あらゆる整数若しくは整数以外の比によるダウンサンプリングが使用されてもよい。

【0049】時間的かつ空間的にスケールされた符号化が例として構成2を使って説明される。ここで、高い空間及び時間解像度を有する入力ビデオシーケンス“in”は、低い空間及び時間解像度を有するベースレイヤーシーケンス及び高い空間及び時間解像度を有する強化レイヤーシーケンスに変換される。これは、信号“in_0”を与えるべく、ブレプロセッサ205におけるダウンサンプル・フィルタリング及びプルダウン処理、並びにミッドプロセッサ215、265におけるアップサンプル・フィルタリング及びポストプロセッサ280におけるノーマル処理を実行することによって表1に示されるように達成される。

【0050】構成3に関して、低若しくは高時間解像度及び高空間解像度を有する入力ビデオシーケンス“in”は、対応する低若しくは高時間解像度及び高空間解像度を有するベースレイヤーシーケンス、並びに対応する低若しくは高時間解像度及び低空間解像度を有する強化レイヤーシーケンスに変換される。これは、強化レイヤーシーケンス“in_1”に対しブレプロセッサ205におけるダウンサンプル・フィルタリング、及びミッドプロセッサ215、265におけるダウンサンプル・フィルタリングを実行することによって、並びに強化レイヤーシーケンス“out_1”に対しポストプロセッサ280におけるアップサンプル・フィルタリングを実行することによって達成される。

【0051】残りの構成は上記例を見れば理解できよう。

【0052】図3は本発明に従う画素のアップサンプリングを示したものである。アップサンプリング・フィルタリングは表1の構成1に関してミッドプロセッサ215、265によって実行される。例えば、CIF解像度(360×288画素)を有するVOPは、2:1アップサンプリングによってITU-R 601解像度(720×576画素)に変換される。CIFイメージの画素310、320、330及び340は、ITU-R

601イメージの画素355、360、365、370、375、380、385及び390を生成するべくサンプリングされる。特に、ITU-R 601画素360は矢印312及び322によって示されるようにCIF画素310及び320をサンプリングすることによって生成される。同様に、ITU-R 601画素365はまた矢印314及び324によって示されるようにCIF画素310及び320をサンプリングすることによって生成される。

【0053】図4はベースレイヤーと強化レイヤー内のVOP間での予測処理の例を示したものである。図2の強化レイヤーエンコーダ210において、強化レイヤーのVOPはP-VOP若しくはB-VOPとして符号化される。この例において、強化レイヤー内のVOPはベースレイヤーVOPよりも高い空間解像度を有し、従ってより大きな領域で描かれている。時間解像度(例えば、フレーム速度)は両レイヤーに対して同一である。該VOPは左から右へ説明順に示されている。

【0054】ベースレイヤーはI-VOP405、B-VOP415及び420並びにP-VOP430を含む。強化レイヤーはP-VOP450及び490並びにB-VOP460及び480を含む。B-VOP415は矢印410及び440に示されるように他のベースレイヤーVOPから予測され、一方、B-VOP420はまた矢印425及び435で示されるように他のベースレイヤーVOPから予測される。P-VOP430は矢印445に示されるようにI-VOP405から予測される。P-VOP450は矢印455に示されるようにベースレイヤーVOPをアップサンプリングすることにより引き出され、一方P-VOP490は矢印495に示されるようにベースレイヤーVOPをアップサンプリングすることによって引き出される。B-VOP460は矢印465及び475で示されるようにベースレイヤーVOPから予測され、またB-VOP480は矢印470及び485で示されるようにベースレイヤーVOPから予測される。

【0055】概して、ベースレイヤー内のI-VOPと時間的に同時(例えば、表示若しくは説明順で)である強化レイヤーVOPはP-VOPとして符号化される。例えば、VOP450はI-VOP405と時間的に同時であり、従ってP-VOPとして符号化される。ベースレイヤー内のP-VOPと時間的に同時である強化レイヤーVOPはP-若しくはB-VOPとして符号化される。例えば、VOP490はP-VOP430と時間的に同時であって、P-VOPとして符号化される。ベースレイヤー内のB-VOPと時間的に同時である強化レイヤーVOPはB-VOPとして符号化される。例えば、B-VOP 460及び480参照。

【0056】I-VOP405及びP-VOP430は強化レイヤーVOP用の予測リファレンスとして使用されるため、それらはアンカーVOPとして知られる。したがって、I-VOP405及びP-VOP430は、強化レイヤー内の対応する予測VOPの符号化前に符号化される。強化レイヤー内のP-VOPの予測リファレンスは、前方(予測)時間リファレンスインジケータであるMPEG-4互換構文のforward_temporal_refiによって特定される。そのようなインジケータはベースレ

イヤー内の時間的に同時のI-VOPを指定する負でない整数である。強化レイヤー内のB-VOPの予測リファレンスはref_select_code, forward_temporal_ref及びbackward_temporal_refによって特定される。以下の表2参照。

ref_select_code	前方時間リファレンスVOP	後方時間リファレンスVOP
00	ベースレイヤー	ベースレイヤー
01	ベースレイヤー	強化レイヤー
10	強化レイヤー	ベースレイヤー
11	強化レイヤー	強化レイヤー

【0057】表2は強化レイヤー内のB-VOP用の予測リファレンス選択を示す。例えば、ベースレイヤー内のI-VOP405及びP-VOP430に対する時間リファレンスコードtemporal_refがそれぞれ0及び3であると仮定する。また、強化レイヤー内のP-VOP450に対するtemporal_refを0とする。その場合、図4において、P-VOP450に対してforward_temporal_ref=0である。矢印465、475、470及び485で示されるB-VOP460及び480の予測リファレンスはそれぞれ、ref_select_code=00、forward_temporal_ref=0及

びback_temporal_ref=3によって特定される。P-VOP490の予測リファレンスはref_select_code=10、forward_temporal_ref=0及びback_temporal_ref=3によって特定される。

【表2】

【0058】ベース及び強化レイヤーの両方の符号化において、予測モードは以下の表3で与えられる2ビットワードのVOP_prediction_typeによって指定される。

【表3】

VOP_prediction_type	Code
I	00
P	01
B	10

【0059】“I”予測タイプはイントラ符号化VOPを示し、“P”予測タイプは予測されたVOPを示し、また“B”予測タイプは双方向予測されたVOPを示す。ベースレイヤーのシーケンス“in_0”に対する符号化処理は、例えば、MPEG-4メインプロファイル若しくはH.263規格に従うノンスケラブル符号化処理と同じである。

【0060】図6は本発明に従う、フレームからフィールドモードへの画素ラインの並べ替え若しくは置換を示す。上記したように、入力VOPがフィールドモードでありかつダウンサンプリングされるとき、生成VOPはフレームモードである。従って、ダウンサンプルされたイメージが空間的にアップサンプリングされるとき、生成VOPはまたフレームモードである。しかし、アップサンプルVOPがアップサンプルVOPから入力VOPを減算することにより差分符号化されるとき、DCTのような空間変換が生成された残差について続けて実行されると当該残差は最適な符号化効率を与えない。すなわち、多くの場合、反対フィールドの画素よりも同一フィールドの画素の間の方がより大きな相関関係が存在するため、残差値の大きさはアップサンプリングされたイメージのラインを偶数及び奇数ラインにグループ化するように置換（すなわち、並べ替え）することによって減少する。

【0061】強化レイヤー内のアップサンプリングされた画素（例えば、ルミナンス）データを表すイメージが符号600で示されている。例えば、イメージ600は8×8ブ

ロックの2：1アップサンプリングによって引き出される16×16マクロブロックであると仮定する。該マクロブロックは偶数番号ライン602、604、606、608、610、612、614、及び616並びに奇数番号ライン603、605、607、609、611、613、615及び617を含む。偶数及び奇数ラインはそれぞれトップ及びボトムフィールドを形成する。マクロブロック600は、領域620とライン602～609の交差によって画成される第1ブロック、領域625とライン602～609の交差によって画成される第2ブロック、領域620とライン610～617の交差によって画成される第3ブロック、及び領域625とライン610～617の交差によって画成される第4ブロックを含む、4つの8×8ルミナンスブロックを有する。

【0062】イメージ600内の画素ラインが、残差を決定しかつDCTを実行する前に本発明に従って同一フィールドルミナンスブロックを形成するべく置換されると、マクロブロック650が形成される。矢印645はライン602～617の並べ替えを示す。例えば、マクロブロック600の第1ラインである偶数ライン602はまたマクロブロック650の第1ラインでもある。偶数ライン604はマクロブロック650内の第2ラインを作る。同様に、偶数ライン606、608、610、612、614及び616はそれぞれマクロブロック650の3番目から8番目のラインを作る。こうして、偶数番号ラインを有する16×8ルミナンス領域680が形成される。第1の8×8ブロックは領域680及び670の交差に

よって画成され、一方第2の8×8ブロックは領域680及び675の交差によって画成される。

【0063】同様に、奇数番号ラインは16×8領域685へ移動させられる。領域685は、領域685及び670の交差によって画成される第1の8×8ブロックから成り、一方第2の8×8ブロックは領域685及び675の交差によって画成される。こうして領域685は、奇数ライン603、605、607、609、611、613、615及び617を有する。

【0064】ここで言及される残差に対して実行されるDCTは、マクロブロック600がマクロブロック650で示されるように並べ替えられるか否かによって、“フィールドDCT”若しくは“フレームDCT”のいずれかである。しかし、本発明はその他の空間変換とともに使用するよう適用することが可能である。フィールドDCTが使用され

$$\sum_{i=0}^6 \sum_{j=0}^{15} |p_{2i,j} - p_{2i+1,j}| + |p_{2i+1,j} - p_{2i+2,j}| > \sum_{i=0}^6 \sum_{j=0}^{15} |p_{2i,j} - p_{2i+2,j}| + |p_{2i+1,j} - p_{2i+3,j}| + \text{bias}$$

【0066】ここで $P_{i,j}$ は、8×8ルミナンスブロックのそれぞれに対しDCTが実行されるすぐ前の空間ルミナンス差（例えば、残差）データである。有利なことに、方程式は第1オーダーの差を使用するため、より単純かつ廉価に実行される。“bias”項は考慮されない非線形効果の説明するファクターである。例えば、bias=64が使用される。もし上記関係が保持されなければ、フレームDCTが使用される。

【0067】上記方程式の左側において、エラー項は反対フィールドの画素差（例えば、偶数から奇数、及び奇数から偶数）に関する。従って、左側は反対フィールドラインのルミナンス値の差の和である。右側において、エラー項は同一フィールドの画素差（例えば、偶数から偶数、及び奇数から奇数）に関する。したがって、右側は同一フィールドラインのルミナンスデータの差の和及びバイアス項である。

【0068】変形的に、絶対値ではなく各エラー項の自乗をとるよう上記方程式を修正することによってフレーム若しくはフィールドDCTのいずれが使用されるべきかを決定するべく、二次オーダーの方程式が使用されてもよい。この場合、“bias”項は必要ない。

【0069】図5は、本発明に従うVOPの空間及び時間スケールを示す。オブジェクトベースのスケラビリティに関して、選択されたVOPのフレーム速度及び空間解像度は強化され、その結果それは残りの領域よりも高い品質を有し、例えば、選択されたオブジェクトのフレーム速度及び／または空間解像度は残りの領域のそれより高い。例えば、ニュースアナウンサーのVOPはスタジオ背景より高い解像度で与えられる。

【0070】軸線505及び506はフレーム番号を示す。ベースレイヤーにおいて、VOP520を含むフレーム510はフレーム0の位置に与えられ、一方VOP532（VOP520に対応する）を有するフレーム530はフレーム3の位置に与えら

るとき、マクロブロックの空間ドメインのルミナンスライン（若しくはルミナンスエラー）は、フレームDCTオリエンテーションからトップ（偶数）及びボトム（奇数）フィールドDCT構成へ置換される。生成マクロブロックは変換され、量子化されかつ普通に可変長符号化される。フィールドDCTマクロブロックが復合化されると、すべてのルミナンスブロックが逆DCT（IDCT）から得られた後に逆置換が実行される。4:2:0クロミナンスデータはこのモードでは有効でない。

【0065】本発明に従ってフィールド若しくはフレームモードDCTを選択するための基準が以下に示される。次の条件を満たすとき、フィールドDCTが選択されなければならない。

【数1】

れる。さらに、矢印512で示されるように、フレーム530はフレーム510から予測される。強化レイヤーはVOP52、524、526及び542を含む。これらのVOPはVOP520及び532に比べ増加した空間解像度を有し、したがってより大きな領域で描かれている。

【0071】P-VOP522は矢印570で示されるようにアップサンプリングVOP520から引き出される。B-VOP524及び526は、矢印572及び576並びに574及び578によりそれぞれ示されるように、ベースレイヤーVOP520及び532から予測される。

【0072】ベース及び強化レイヤーシーケンスを作るために使用される入力ビデオシーケンスは完全な解像度（例えば、National Television Standards Committee (NTSC)に対応するITU-R 601用の720×480またはPhase A Iteration Line (PAL)に対応するITU-R用の720×576）及び完全フレーム速度（NTSCに対応するITU-R用30フレーム/60フィールド若しくはPALに対応するITU-R601用25フレーム/50フィールド）を有する。スケラブル符号化が実行され、その結果オブジェクトの解像度及びフレーム速度は強化レイヤー符号化を使って保持される。VOP520及び532から成るベースレイヤー内のビデオオブジェクトはより低い解像度（例えば、完全解像度VOPの4分の1サイズ）及びより小さいフレーム速度（例えば、オリジナルフレーム速度の3分の1）を有する。さらに、強化レイヤーにおいて、VOP520のみが強化される。フレーム510の残りは強化されない。たったひとつのVOPのみが示されているが、仮想的にあらゆる数のVOPが与えられ得る。さらに、2つ以上のVOPが与えられるとき、すべて若しくは選択されたもののみが強化されてもよい。

【0073】ベースレイヤーシーケンスはオリジナルシーケンスをダウンサンプリングしかつフレームドロップングすることによって生成される。その後、ベースレイ

ヤーVOPはプログレッシブ符号化ツールを使ってI-VOP若しくはP-VOPとして符号化される。入力ビデオシーケンスがインターレースされると、入力されたインターレースビデオシーケンスはプログレッシブビデオシーケンスを生成するため、フィールド／フレーム移動推定及び補償のようなインターレース符号化ツール、並びにフィールド／フレームDCTは使用されない。強化レイヤーVOPは時間的及び空間的スケラブルツールを使って符号化される。例えば、強化レイヤーにおいて、VOP522及びVOP542は空間スケラビリティを使ってP-VOPとして符号化される。VOP524及びVOP526は、時間的スケラブルツールを使って、ベースレイヤーリファレンスVOP、すなわちVOP520及びVOP532のそれぞれアップサンプリングされたVOPからのB-VOPとして符号化される。

【0074】本発明の更なる態様として、移動ベクトル調査範囲を縮小することによって、B-VOPの移動推定に対する符号化の複雑さを減少させる技術が開示される。当該技術はフレームモード及びフィールドモードの両方の入力ビデオシーケンスに応用可能である。特に、リファレンスVOPの調査中心は、リファレンスVOPにおいて独立の網羅的調査を実行するのではなく、対応するベースレイヤーVOPの移動ベクトルをスケラリングすることによって決定される。典型的にそのような網羅的調査は例えば、水平方向に ± 64 画素で垂直方向に ± 48 画素の範囲をカバーするため、開示された技術よりも効率が悪い。

【0075】強化レイヤー内のB-VOP524及び526の移動ベクトルに対する調査中心は、以下のように決定され、 $MV_f = (m/n \cdot TR_b \cdot MV_p) / TR_p$ 、 $MV_b = (m/n \cdot (TR_b - TR_p) \cdot MV_p) / TR_p$ ここで、 MV_f は前方移動ベクトル、 MV_b は後方移動ベクトル、 MV_p はベースレイヤー内のP-VOP（例えば、VOP532）に対する移動ベクトルであり、 TR_b は過去のリファレンスVOP（例えば、VOP520）と強化レイヤー内の現B-VOPとの間の時間差であり、 TR_p は過去のリファレンスVOPとベースレイヤー内の未来のリファレンスP-VOP（例えば、VOP532）との間の時間差である。 m/n は、強化レイヤーVOPの空間解像度に対するベースレイヤーVOPの空間解像度の比である。すなわち、ベースレイヤーVOP若しくは強化レイヤー内のB-VOPのいずれかが、入力ビデオシーケンスに対して比 m/n でダウンサンプリングされてもよい。図5の実施例において、 m/n は、強化レイヤーVOPを与えるべく続けてアップサンプリングされるベースレイヤーVOPのダウンサンプリング比である。 m/n は1より小さいか、1と等しいか、または1より大きい。例えば、B-VOP524、 $TR_b=1$ 、 $TR_p=3$ 及び2:1ダウンサンプリング（すなわち、 $m/n=2$ ）に対して、 $MV_f=2/3MV_p$ 及び $MV_b=-4/3MV_p$ を得る。すべての移動ベクトルは2次元であることに注意すべきである。該移動ベクトルの調査範囲は、例えば、その中心が MV_f 及び MV_b によって決定されるとき

の、 16×16 の矩形範囲である。移動ベクトルは移送データストリーム内で強化レイヤー及びベースレイヤービデオデータとともに通信され、ビデオデータを復元化する際にデコーダによって回復される。

【0076】概して、本発明に従うインターレースされたビデオの符号化に対して、より良い性能を発揮するためにインターレース符号化ツールが使用される。これらのツールは、イントラ-マクロブロック及びインター-ディファレンス・マクロブロックに対するフィールド／フレームDCT、並びに、トップフィールドからボトムフィールドへ、トップフィールドからトップフィールドへ、ボトムフィールドからトップフィールドへ及びボトムフィールドからボトムフィールドへのフィールド予測を含む。

【0077】上記表1に記載された構成に対し、これらインターレース符号化ツールは以下のように組み合わせられる。

【0078】(1)両方のレイヤーに対し低い空間解像度を有する構成に対して、プログレッシブ（フレームモード）符号化ツールのみが使用される。この場合、2つのレイヤーは例えば、立体ビデオ信号内の異なるビューシーケンスを符号化する。立体ビデオ符号化に対し、右像（強化レイヤー）シーケンスに対する移動推定調査範囲は 8×8 画素である。この 8×8 （完全画素）調査範囲は対応するVOPのベースレイヤー内の対応するマクロブロックの同じタイプの移動ベクトルに中心が置かれる。

【0079】(2)ベースレイヤー内で低い空間解像度を有しかつ強化レイヤー内で高い空間解像度を有する構成に対して、インターレース符号化ツールは強化レイヤーシーケンスに対してのみ使用される。強化レイヤーシーケンスを符号化するための移動推定調査範囲は 8×8 （完全画素）である。この 8×8 の調査範囲は、対応するVOPのベースレイヤー内の対応するマクロブロックの再スケラリングされた（すなわち2のファクタ）同じタイプの移動ベクトルに中心が置かれる。フィールドベースの推定及び予測は、強化レイヤーの調査及び補償の際にのみ使用される。

【0080】(3)ベースレイヤー内で高い空間解像度を有し、強化レイヤー内で低い空間解像度を有する構成に対して、インターレース符号化ツールはメインレベルでのMPEG-2メインプロファイルと同様にベースレイヤーシーケンスに対してのみ使用される。強化レイヤーシーケンスを符号化するための移動推定調査範囲は 4×4 （完全画素）である。この 4×4 調査は、対応するVOPのベースレイヤー内の再スケラリングされた（すなわち、 $1/2$ のファクタ）同じタイプの移動ベクトルに中心が置かれる。上記表1の構成2に対して、例えば、2つのレイヤーのシーケンスの符号化は異なる時間単位速度を有する。

【0081】図7は、本発明に従う空間及び時間スケラリングを有するピクチャー・イン・ピクチャー（PIP）若

しくはプレビュー・チャンネル・アクセス応用を示した図である。PIPに関して、二次番組がテレビ上で観られるメイン番組のサブセットとして与えられる。該二次番組はより小さい領域を有するため、視聴者は解像度の低いイメージを見分けることになり、その結果PIPイメージの時間及び/または空間解像度は帯域幅を確保するべく減少させられる。

【0082】同様に、プレビュー・アクセス・チャンネル番組は視聴者に対して、料金を購入できる番組の無料低解像度サンプルを与えることができる。この応用はプレビュー用に許可チャンネル（例えば、ペイ・パー・ビュー（pay-per-view））の数分の無料アクセスを与える。プレビュー・アクセス・チャンネル内で符号化されたビデオは、より低い解像度及びより小さいフレーム速度を有する。デコーダはそのようなプレビューチャンネルに対してアクセス時間を制御する。

【0083】上記表1における時間-空間スケーラブル符号化の構成2が、ベースレイヤー及び強化レイヤーの両方の復合化からの出力より低い空間解像度を有するベースレイヤーの復合化からの出力を与えるべく使用されてもよい。ベースレイヤー内のビデオシーケンスは小さいフレーム速度で符号化され、一方強化レイヤーはより高いフレーム速度で符号化される。

【0084】例えば、ベースレイヤー内のビデオシーケンスはCIF解像度及び15フレーム/秒の速度を有し、一方対応する強化レイヤー内のビデオシーケンスはITU-R601解像度及び30フレーム/秒のフレーム速度を有する。この場合、強化レイヤーはNTSCビデオ規格に一致し、一方CIF規格に一致するPIP若しくはプレビューアクセス機能がベースレイヤーによって与えられる。したがって、PIP機能は、メインレベル規格でのMPEG-2メインプロファイルと同様の符号化効率及び複雑性を有するスケーラブル符号化によって与えられる。

【0085】ベースレイヤーは低い空間解像度VOP705及び730を含む。さらに、ベースレイヤーの時間的解像度は強化レイヤーのその3分の1である。該強化レイヤーは高い空間解像度VOP750、760、780及び790を含む。P-VOP750は矢印755で示されるように、I-VOP705をアップサンプリングすることによって引き出される。B-VOP760は矢印765及び775によって示されるようにベースレイヤーVOPから予測される。B-VOP780は矢印770及び785によって示されるようにベースレイヤーVOPから予測される。P-VOP790は矢印795で示されるようにP-VOP730をアップサンプリングすることによって引き出される。

【0086】図8は本発明に従う立体ビデオ応用を示す。立体ビデオ機能は、書類ISO/IEC JTC1/SC29/WG11 N1196で説明されるMPEG-2 マルチビュープロファイル(MVP)システム内で与えられる。ベースレイヤーは左像に割り当てられ、強化レイヤーは右像に割り当てられる。

【0087】符号化効率を改良するために、強化レイヤ

ー画像はベースレイヤーよりも低い解像度で符号化される。例えば、上記表1の構成4は、ベースレイヤーがITU-R601空間解像度を有し、一方強化レイヤーがCIF空間解像度を有するところで使用される。強化レイヤー画像の予測用のベースレイヤーのリファレンス画像がダウンサンプリングされる。したがって、強化レイヤー画像用のデコーダはアップサンプリング処理を含む。したがって、適応フレーム/フィールドDCT符号化は強化レイヤーではなくベースレイヤーで使用される。

【0088】ベースレイヤーはVOP805、815、820及び830を含み、一方強化レイヤーはVOP850、860、880及び890を含む。B-VOP815及び820は矢印810、840及び835、825によってそれぞれ示されるように、他のベースレイヤーVOPを使って予測される。P-VOP830は矢印845で示されるようにI-VOP805から予測される。P-VOP850は矢印855で示されるようにI-VOP805をダウンサンプリングすることによって引き出される。B-VOP860は矢印865及び875によって示されるように、ベースレイヤーVOPから予測される。B-VOP880は矢印870及び885によって示されるように、ベースレイヤーVOPから予測される。P-VOP890は、矢印895で示されるようにP-VOP830をダウンサンプリングすることによって引き出される。

【0089】変形的に、同じ空間解像度及びフレーム速度を有するベース及び強化レイヤーに対して、上記表1の構成7が使用される。この場合、ベースレイヤーの符号化処理はMPEG-4 VM 非スケーラブル符号化またはメインレベル標準でのMPEG-2メインプロファイル内で説明されるような非スケーラブル符号化処理と同じであり、適応フレーム/フィールドDCT符号化は強化レイヤー内で使用される。

【0090】本発明の他の応用において、非同期転送モード(ATM)通信技術が与えられる。概して、ATMネットワークを通じたビデオ信号送信へのトレンドは急速に高まっている。これはこれらのネットワークの可変ビット速度(VBR)特性によるものであり、それは一定ビット速度(CBR)送信を超えたいくつかの利点を有する。例えば、VBRチャンネルにおいて、ほぼ一定の画像品質が達成され得る。さらに、ATMネットワーク内のビデオソースは、CBRシステム内の弾性バッファリングによって、ビデオ信号データ速度の長期間平均が短期間平均よりも小さいために、CBRチャンネルを通じて送信するよりも小さい送信ビット速度を要求しながら、統計的に多重化される。

【0091】しかし、ATMネットワークの利点にも関わらず、それらは混雑という大きな欠点を被る。混雑したネットワークにおいて、ビデオパケットは出カールトを見つるべく待機(queued)させられる。長く遅延したパケットは到着が遅すぎてレシーバでは使用されず、その結果デコーダによって廃棄される。したがって、ビデオcodecはパケット損失に抵抗するよう設計されなければならない。

【００９２】ビデオコーダーにパケット損失の免疫を与えるため、本発明の時間的-空間的スケーラブル符号化技術が使用される。特に、ベースレイヤーからのビデオデータは高い優先順位で送信され、ATMネットワークの保証されたビット速度に調節される。もしチャネルが保証されないために混雑が生じれば、強化レイヤーからのビデオデータパケットは失われる。もし強化レイヤーパケットが受信されると、画像品質は改良される。上記表１の構成１を使用する符号化手法がこの結果を達成するのに使用され得る。この手法は、ベースレイヤーが高優先順位のレイヤーであるところの予測モードとの関係ですで説明したように、図４で示されるようにして達成される。したがって、優先順位がより高く、ビット速度のより小さいデータはベースレイヤー内で通信され、優先順位がより低く、ビット速度がより高いデータは強化レイヤー内で通信される。

【００９３】同様にして、そのようなスケーラブル符号化はまた、ビデオ符号化の際並びにインターネット、イントラネット及び他の通信ネットワークを通じて送信する際に使用される。

【００９４】したがって、本発明は、デジタルビデオシーケンス内にビデオオブジェクト平面(VOP)を含むビデオイメージの時間的及び空間的スケーリングを与えるための方法及び装置を与えることがわかる。本発明のひとつの態様において、符号化効率はスケーリングされたフィールドモード入力ビデオシーケンスを適宜圧縮することによって改良される。強化レイヤー内のアップサンプリングされたVOPは、線形基準に基づいたオリジナルのビデオシーケンスに対しより大きな相関関係を与えるべく並べ替えられる。生成された残差はDCTのような空間変換を使って符号化される。本発明の他の態様において、移動補償手法は、ベースレイヤーVOPに対してすでに決定された移動ベクトルをスケーリングすることによって強化レイヤーVOPを符号化するために与えられる。スケーリングされた移動ベクトルによってその中心が定義されるところの縮小された調査範囲が画成される。その技術はスケーリングされたフレームモード若しくはフィールドモードの入力ビデオシーケンスとともに使用するのに適している。

【００９５】付加的に、特定のスケーラブルな符号化結果を与えるべくさまざまなcodecプロセッサ構成が与えられた。立体ビデオ、ピクチャー・イン・ピクチャー、プレビュー・アクセス・チャネル、及びATM通信を含むスケーラブル符号化の応用もまた議論された。

【００９６】発明はさまざまな特定の実施例との関連で説明されてきたが、請求の範囲に記載された発明の思想及び態様から離れることなくさまざまな付加及び修正が

可能であることは当業者の知るところである。例えば、２つのスケーラビリティレイヤーが議論されてきたが、２つ以上のレイヤーが与えられてもよい。さらに、単純化のために矩形若しくは正方形のVOPがいくつかの図面中で与えられたが、本発明は任意の形状のVOPでの使用に同じように適している。

【図面の簡単な説明】

【図１】図１は、本発明に従うビデオオブジェクト平面(VOP)の符号化及び復合化処理を示したものである。

【図２】図２は、本発明に従うVOPコーダ及びデコーダのブロック図である。

【図３】図３は、本発明に従う画素のアップサンプリングを示したものである。

【図４】図４は、ベースレイヤーと強化レイヤーのVOPの間の予測処理を例示したものである。

【図５】図５は、本発明に従うVOPの空間的及び時間的スケーリングを示したものである。

【図６】図６は、本発明に従って画素ラインをフレームからフィールドモードへ並べ替えた様子を示す。

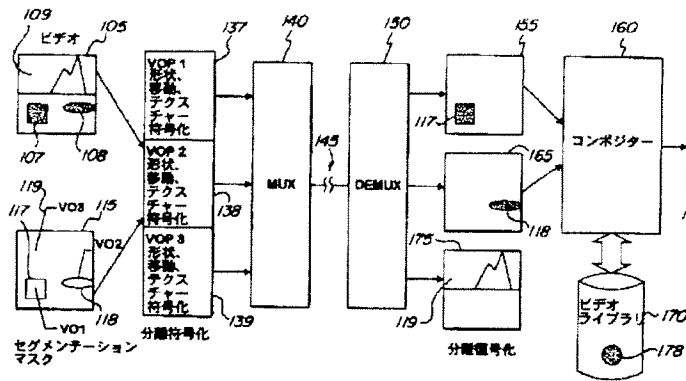
【図７】図７は、本発明に従う空間的及び時間的スケーリングを有する、ピクチャー・イン・ピクチャー(PIP)若しくはプレビュー・チャネル・アクセスアプリケーションを示す。

【図８】図８は、本発明に従う立体ビデオアプリケーションを示す。

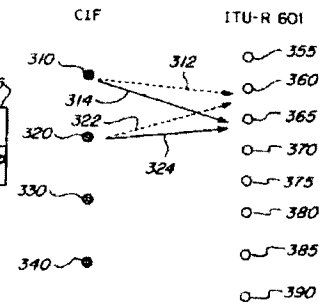
【符号の説明】

105	フレーム
107	前景エレメント
108	前景エレメント
109	背景エレメント
115	フレーム
117	ビデオオブジェクト平面
118	ビデオオブジェクト平面
119	ビデオオブジェクト平面
137	エンコーダ
138	エンコーダ
139	エンコーダ
140	マルチプレкса
145	チャネル
150	デマルチプレкса
155	フレーム
160	コンポジター
165	フレーム
170	ビデオライブラリ
175	フレーム
178	ビデオオブジェクト平面
185	フレーム

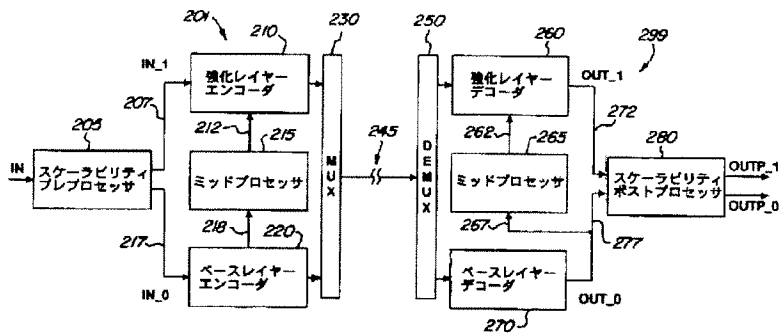
【図 1】



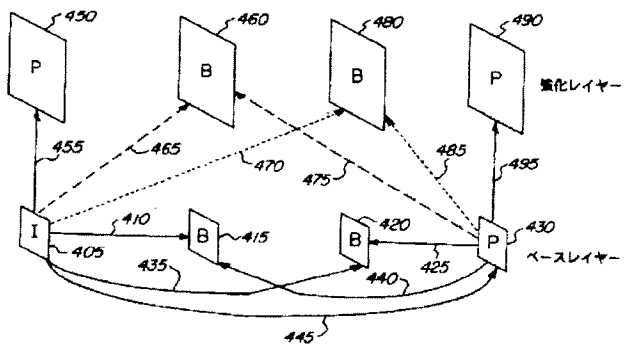
【図 3】



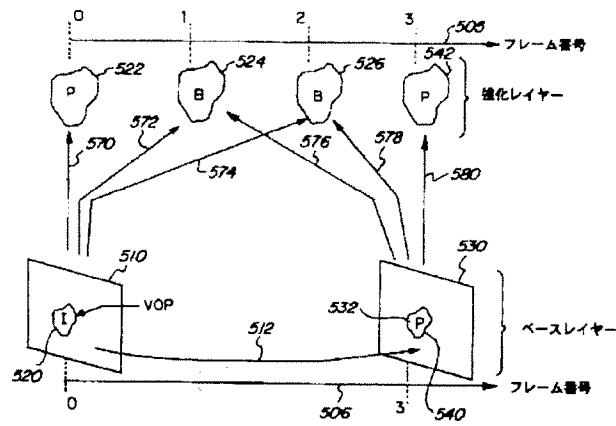
【図 2】



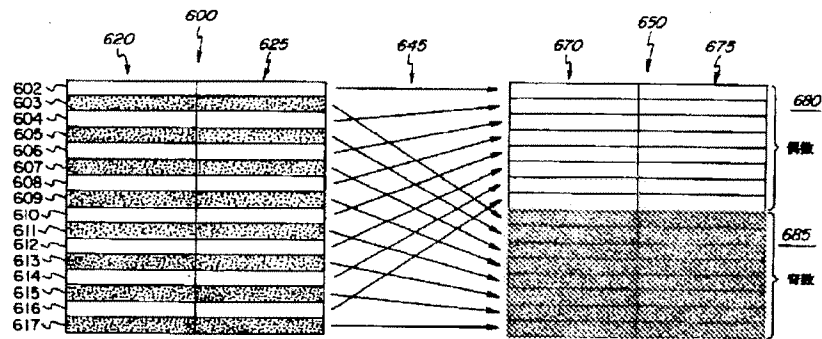
【図 4】



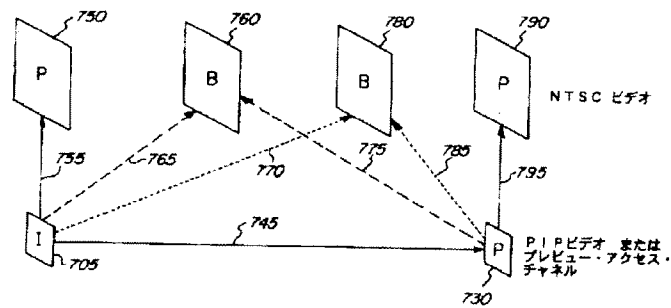
【図5】



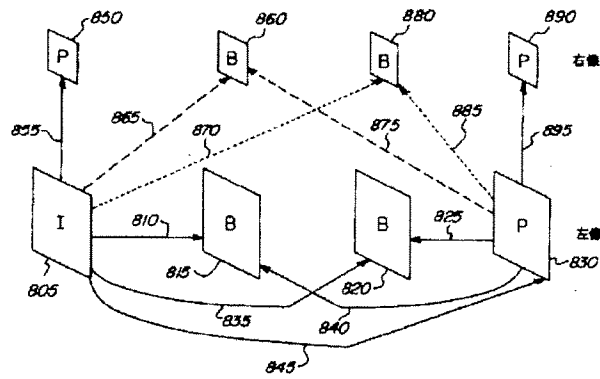
【図6】



【図7】



【図 8】



フロントページの続き

(71)出願人 598045380

101 Tournament Drive
Horsham, Pennsylvania, The United States of America

(72)発明者 ガネッシュ・ラジャン

アメリカ合衆国カリフォルニア州サンディエゴ, ショアライン・ドライブ 7150 ナンバー3202

(72)発明者 マンダヤム・ナラシムハン

アメリカ合衆国カリフォルニア州サンディエゴ, ジャネット・プレイス 13134